

پردازش تصویر با OPENCV

استفاده از Template Matching

در این روش بخشی از فریم که می‌خواهیم دنبال شود راه به صورت عکس به الگوریتم می‌دهیم و سپس الگوریتم با چک کردن پیکسل به پیکسل عکس آن را پیدا کرده و به صورت یک نمودار حرارتی به صورتی که قسمتی که عکس داده شده وجود دارد به رنگ روشن در می‌آید، نشان داده می‌شود.

مزایا

۱. راحتی در برنامه نویسی
۲. خروجی با استفاده از نمودار حرارتی

معایب

۱. عدم کاربرد واقعی
۲. عدم تشخیص صحیح در صورتی که عکس زاویه دار و یا برعکس باشد

استفاده از Watershed Algorithm

مطالعه پردازش تصویر، یک تبدیل تعریف شده بر تصاویر سیاه سفید است. این نام به صورت استعاری به یک حوضه آبریز زمین‌شناسی یا تقسیم زهکشی که حوضه‌های زهکشی مجاور را جدا می‌کند، اشاره دارد. تبدیل آب‌پخش‌شان یا کار روی نقشه توپوگرافی تصویر عمل می‌کند، ارتفاع آن را با روشنایی هر نقطه نشان می‌دهد و خطوطی را که در امتداد تپه‌ها قرار دارند پیدا می‌کند. [اطلاعات بیشتر](#)

عمومی‌ترین کاربرد این الگوریتم برای جدا کردن تو بخش که به یک دیگر متصل هستند، می‌باشد

استفاده از Feature Detection

در این روش با دادن یک عکس پایه الگوریتم ویژگی‌های آن را بدست آورده و در عکس اصل به دنبال این ویژگی‌ها می‌باشد

دارای دو الگوریتم OBV و SIFT می‌باشد که الگوریتم دوم SIFT دارای دقت بیشتری می‌باشد

مزایا:

۱. راحتی در کار
۲. سرعت پردازش بالا
۳. RealTime

معایب:

۱. دقت پایین
۲. خطای بسیار
۳. عدم مشخص کردن مکان شی
۴. عدم مشخص کردن نقاط X,Y
۵. عدم مشخص کردن زاویه

استفاده از Haar Cascade

فایل‌های Haar Cascade فایل‌های XML می‌باشند که به صورت کلی یک Image Classifier می‌باشد که برای تشخیص یک شی در تصویر به کار می‌روند

مزایا

۱. کد نویسی کم و راحت

۲. کم حجم بودن فایل Train شده
 ۳. امکان ردیابی و تشخیص
 ۴. امکان تشخیص چند شی
 ۵. استفاده گسترده در Face Detection و ...
- معایب
۱. تأخیر بالا در انتقال تصویر
 ۲. سخت بودن ساخت فایل ها
 ۳. در صورت تهیه فایل با نمونه های کم و رزولوشن پایین دارای دقت بسیار کمی میباشد

استفاده از Opencv Tracker

کتاب خانه Opencv دارای API های قدرتمندی برای دنبال کردن شی میباشد که میتوان از :

۱. BOOSTING
 ۲. MIL
 ۳. KCF
 ۴. TLD
 ۵. MEDIANFLOW
- نام برد که بهترین آن ها در وضعیتی که تست شد Boosting بود
- مزایا :
۱. راحتی در کار و استفاده
 ۲. سرعت بالا
 ۳. دقت بالا
 ۴. استفاده از الگوریتم های مختلف در شرایط مختلف
 ۵. ردیابی کامل شی
 ۶. به دست آوردن کامل نقاط X,Y
- معایب
1. عدم ردیابی دو شی

استفاده از Shape Detection Object

این الگوریتم تهیه شده از چند الگوریتم مختلف بوده و توسط خودمان ترکیب و تهیه شده که با استفاده از ترکیب رنگ های [HSV](#) کار میکند که بعد از اضافه کردن چند فیلتر مانند GaussianBlur تبدیل می شود سپس لبه ها و شی توسط find Contours پیدا شده بزرگ ترین لبه انتخاب شده و نشان داده می شود

- مزایا
1. پیدا کردن شی
 2. ردیابی شی
 3. مشخص کردن شی
 4. دسترسی به نقاط X,Y
 5. دسترسی به زاویه شی
 6. امکان تنظیم رنگ ها برای بهتر پیدا کردن شی ها
- معایب
1. این الگوریتم فقط وجود اشیاء رو مشخص میکند نه نوع آن

معرفی دیتاست ها

دیتاست ها مجموعه عظیم از عکس ها در یک حوزه خاص آماده برای استفاده در شبکه عصبی مصنوعی میباشد.

- **MNIST**: دیتاست ارقام دست نویس انگلیسی (خیلی معروف و پرکاربرد)
- **دیتاست هدا**: دیتاست ارقام دست نویس فارسی (۶۰ هزار تصویر)
- **NIST**: مثل MNIST اما بزرگتر.
- **Perturbed NIST**: دیتاستی که در کلاس Yoshua ایجاد شد (همون دیتاست NIST هست به اضافه مقدار زیادی deformation یا تغییر شکل (یعنی تصاویر همون دیتاست رو با تغییرات مختلف ذخیره کرده و دیتاست بزرگتری رو ایجاد کرد)
- **CIFAR۱۰ / CIFAR۱۰۰**: این دوتا دیتاست در اصل یکی هستند و تفاوت اونها در تعداد دسته های اونهاست. این دیتاست شامل تصاویر رنگی با اندازه ۳۲*۳۲ هست و شامل ۶۰ هزار تصویر میشه.
- **Caltech ۱۰۱**: دیتاستی شامل تصاویری از اشیاء در ۱۰۰ دسته مختلف ()
- **Caltech ۲۵۶**: این دیتاست هم شامل تصاویری از اشیاء مختلف در ۲۵۶ دسته هست
- **Caltech Silhouettes**: این دیتاست همونطور که از اسمش مشخصه حاوی تصاویر دودویی شده دیتاست Caltech ۱۰۱ با اندازه ۲۸*۲۸ هست
- **STL-۱۰**: این دیتاست هم مثل CIFAR۱۰ برای استفاده در زمینه های unsupervised feature learning, deep learning, self taught learning مورد استفاده قرار میگیره, فرقی با CIFAR۱۰ در اینه که تغییرات (modification) های بیشتری داره.
- **The Street View House Numbers (SVHN) Dataset**: این دیتاست هم شماره پلاک خونه ها هست.
- **NORB**: تصاویر گرفته شده با دوربین از اسباب بازیها تحت حالتها و شدت نورهای مختلف. (منظورم اینجور تصاویر هست **کلیک کنید تا به نمونه رو ببینید**)
- **Imagenet**: ایمیج نت در اصل اسم یک رقابت تو حوزه تشخیص تصاویر هست که دیتاست های مختلفی داره. هر سال هم برگزار میشه و در اصل تو رقابت سال ۲۰۱۲ اون بود که الکس کریژوسکی و جف هینتون تونستن با دیپ لرنینگ و استفاده از شبکه کانولوشن مقام اول رو بدست بیارن و دیپ لرنینگ اینطور متولد بشه! در حال حاضر معروف ترین و پر استفاده ترین دیتاست تصاویر طبیعی که بیش از یک میلیون تصویر در ۱۰۰۰ دسته داره و اندازه تصاویر ۲۲۷ در ۲۲۷ شروع میشه مربوط به دیتاست سال ۲۰۱۲ اون هست. انواع مختلفی از دیتاستها هست که از نظر حجمی و تعداد و... متفاوت هستند. حجم این دیتاست (سال ۲۰۱۲) بیش از ۱۵۰ گیگابایت هست. ورژن ۲۰۱۵ اون هم با احتساب همه زمینه ها بیش از ۱ ترابایت هست. (بعدا در این مورد بیشتر توضیح میدم)
- **Pascal VOC**: این هم مثل ایمیج نت یه رقابت تو حوزه کامپیوتر ویژن هست که دیتاست های مختلفی داره.
- **Labelme**: یه دیتاست بزرگ از تصاویر حاشیه نویسی شده (اصطلاحا میگوین annotated images که خیلی کاربردی هست) (توضیحات بعدا!!)
- **COIL ۲۰**: اشیاء مختلف در ۲۰ دسته که در تمامی زوایا ۳۶۰ درجه عکس برداری شده ازشون
- **COIL۱۰۰**: مثل بالایی با این تفاوت که اینجا ۱۰۰ دسته وجود داره

- **Arcade Universe** - یه جنریتور (تولید کننده) دیتاست مصنوعی هست که با استفاده از تصاویر شامل sprite های بازی هایی مثل تتریس پنتومینو و تترومینو تصویر جدید تولید میکنه. این جنریتور بر اساس نمونه ایجاد شده توسط o.Breleux بنام bugland ایجاد شده.
- در زیر هم مجموعه ای از دیتاست هایی که با ایده از **BabyAISchool** ایجاد شدند میبینید :
- **BabyAIShapesDatasets** : تمایز بین سه شکل ساده
- **BabyAllImageAndQuestionDatasets** : یه دیتاست سوال-تصویر-جواب
- **(DeepVsShallowComparisonICML۲۰۰۷)** : دیتاست هایی که منحصر برای ارزیابی معماری های عمیق ایجاد شدند:
- **MnistVariations** : این دیتاست شامل تغییرات کنترل شده در دیتاست MNIST هست
- **RectanglesData** : تمایز بین مستطیل های عریض و بلند
- **ConvexNonConvex** : تمایز بین اشکال convex و غیر convex (همون اشکال کاسه شکل(محدب) و غیر کاسه شکل(غیرمحدب)!!)
- **BackgroundCorrelation** : دیتاست شامل درجه کنترل شده ارتباط (correlation) در پس زمینه های نویزی MNIST هست.

دیتاست های تصاویر صورت

- **Labelled Faces in the Wild** : یکی از دیتاست های پرکاربرد که شامل ۱۳ هزار تصویر از افراد مختلف هست که از اینترنت (وب) جمع آوری شده و با نام هر فرد برچسب گذاری شده
- **Toronto Face Dataset** : دیتاست تصاویر صورت دانشگاه تورونتو
- **Olivetti** : تعدادی تصویر از افراد مختلف
- **Multi-Pie** : دیتابیس صورت مالتی پای یا The CMU Multi-PIE Face Database
- **Face-in-Action**
- **JACFEE** : تصاویر صورت افراد ژاپنی و قفقازی به همراه حالات مختلف احساسی
- **FERET** : دیتابیس فناوری تشخیص صورت یا همون The Facial Recognition Technology Database
- **mmifacedb** : دیتابیس حالات صورت MMI یا همون MMI Facial Expression Database
- **IndianFaceDatabase** : دیتابیس از تصاویر صورت افراد هندی !

• [Yale Face Database B](#) و [Yale Face Database](#)

متن

• [newsgroups](#) ۲۰: عمل دسته بندی، نگاشت تکرر کلمات به شناسه newsgroup

• [Reuters \(RCV*\) Corpora](#): پیش بینی متن/موضوع یا text/topic prediction

• [Penn Treebank](#): برای پیش بینی کلمه یا کاراکتر بعدی مورد استفاده قرار میگیره.

• [Broadcast News](#): یه دیتاست بزرگ متنی که بطور کلاسیک برای پیش بینی کلمه بعدی مورد استفاده قرار میگیرفت.

• دیتاست ویکی پدیا!

• [Multidomain sentiment analysis dataset](#): دیتاست تحلیل نیت چند دامنه ای

صدا

• [TIMIT Speech Corpus](#): دسته بندی واج ها (اصوات) یا همون phoneme classification

• [Aurora](#): تیممیت (Timit) به همراه نويز و اطلاعات اضافی

Symbolic Music Datasets

• [Piano-midi.de](#): قطعات کلاسیک پیانو

• [Nottingham](#): بیش از هزار فولک تون

• [MuseData](#): کتابخانه الکترونیکی از امتیازات موسیقی کلاسیک

• [JSB Chorales](#): مجموعه ای از کورالهای ۴ بخشی هارمونیک

Recommendation Systems

• [MovieLens](#): دتا دیتاست از اینجا قابل دریافت هست. اولین دیتاست شامل ۱۰۰ هزار رتبه بندی برای ۱۶۸۲ فیلم توسط ۹۴۳ کاربر هست. که به ۵ بخش مختلف تقسیم شده. دومین دیتاست حدود ۱ میلیون رتبه بندی (ریتینگ) برای ۳۹۰۰ فیلم هست که توسط ۶۰۴۰ کاربر داده شده.

• [Jester](#): این دیتاست شامل ۴۱ میلیون رتبه بندی پیوسته (از -۱۰۰۰۰ تا +۱۰۰۰۰) از ۱۰۰ تا ۷۳۴۲۱ کاربر داده شده.

• [Netflix Prize](#): نت فلیکس یک ورژن از دیتاست رتبه بندی فیلمهای رو ارایه کرده که شامل ۱۰۰ میلیون رتبه بندی هست که توسط ۴۸۰ هزار کاربر انجام شده که بین ۱ تا تمام ۱۷۷۷۰ فیلم را رتبه بندی کردند.

- **Book-Crossing dataset**: این دیتاست از انجمن Book-Crossing هست که شامل ۲۷۸۸۵۸ کاربره که ۱۱۴۹۷۸۰ امتیاز رو در مورد ۲۷۱۳۷۹ کتاب دادند .

متفرقه

-
- دیتاست Musk
 - **CMU** دیتابیس Motion capture هست !
 - **Brodatz dataset**: یه دیتاست برای texture modeling هست!
 - **Million Song dataset**: یک دیتاست شامل ۱ میلیون آهنگ!
 - **Merck Molecular Activity** رقابت فعالیت مولکولی مرک!
 - دیتاست های مختلف در زمینه Sentiment analysis
-

دیتاست های پزشکی

نکته : اگر دیتاستی رو اینجا پیدا نکردید از مخازن دیتاست که در بالا لینکشون قرار داده شده استفاده کنید.

- **Pulmonary hypertension Datasets**: این سایت دارای دسته های مختلف برای انواع مختلفی از سرطان هاست (هم انسان و هم حیوان و...) و دیتاست بسیار بزرگی رو شامل میشه (بیش از هزاران نمونه)
- **Visual Concept Extraction Challenge in Radiology** (VISCERAL): تصاویر رادیولوژی از ساختارهای آناتومیکال (مثل ششها ، کلیه ها و...) به فرمتهای مختلف (CT یا MR) که بصورت دستی حاشیه نویسی شده اند.
- **The Cancer Imaging Archive**: دیتاست های مختلف از انواع سرطان ها (سرطان شش ، myeloma ، carcinoma و...) با فرمت های مختلف
- **Grand Challenges in Biomedical Image Analysis**: مجموعه ای از رقابتهای بیومدیکال که به منظور تسهیل قیاس بین راه حل ها و روشهای موجود و جدید ایجاد شده . در حال حاضر (زمان نگارش این مطلب) ۹۲ رقابت وجود دره . هر رقابت دارای دیتاست خاص خود میباشد.
- **The Lung Image Database Consortium** : مجموعه تصاویر (LIDC-IDRI) که . یک مجموعه از تصاویر سی تی اسکن از سرطان شش با ضایعات (lesions) حاشیه نویسی شده
- **Kaggle diabetic retinopathy**: تصاویر شبکیه چشم (retinal) با کیفیت بالا که به منظور تشخیص retinopathy دیابتی مورد استفاده قرار میگیرند و از مقیاس شدت ۰ تا ۴ توسط تکنسین ها حاشیه نویسی شده اند.
- **International Symposium on Biomedical Imaging**: (۲۰۱۵) رقابت بزرگ بین المللی که دارای دیتاست های خاص خود هستند.
- **Multiple sclerosis lesion segmentation** : مجموعه ای از تصاویر MRI مغز برای تشخیص ضایعات مربوط به بیماری ام اس (MS lesions).

• **Multimodal Brain Tumor Segmentation Challenge**: | (BRATS) دیتاست بزرگی از اسکنهای رزونانس مغناطیسی تومور مغزی (brain tumor magnetic resonance scan) این دیتاست از سال ۲۰۱۲ در حال گسترش هست (هر ساله!)

• **Coding+Cancer**: تصاویر ماموگرافی دیجیتا و تشخیص سرطان شش

دیتاست های متفرقه (بدون ترتیب)
مجموعه ای از آدرس سایت های دیتابیس و دیتاست : (از گروه تلگرام دیپ لرنینگ (لینک در بخش منابع یادگیری اومده))

- انواع DataSet ها
- دیتاست مربوط به یوتیوب :
- دیتاست مربوط به داده های شبکه
- مجموعه داده های طبقه بندی شده در دیتاپول در موضوعات مختلف
- دیتاست مربوط به تشخیص چهره
- دیتاست مربوط به دسته بندی حملات شبکه
- دیتاست مربوط به وام های داده شده ، دیر کرد و غیره به مشتریان است
- دیتاست مربوط به تصادفات جاده ای
- دیتاست مربوط به گربه ها
- دیتاست مربوط به حراجی آنلاین ebay :
- دیتاست مربوط به موسیقی و صدا
- دیتاست مربوط به انواع شبکه های موجود و وب سایت ها - دانشگاه استنفورد
- دیتاست های مربوط به یادگیری ماشین
- گروه داده کاوی دانشگاه تگزاس
- وب سایت های مجموعه داده:
- اطلاعات مجموعه سایت های استک اورفلو (StackExchange)
- <http://stackoverflow.com/questions/۳۸۱۸۰۶/large-public-datasets>
- <https://networkdata.ics.uci.edu/resources.php>
- <https://www.kaggle.com>
- <http://www.kdnuggets.com/datasets/index.html>

- <http://socialcomputing.asu.edu/pages/datasets>
- <http://nodexlgraphgallery.org/Pages/Default.aspx>
- <http://konect.uni-koblenz.de/networks>

بخش دوم (نا مرتب)

1. [Google House Numbers](#) from street view
2. [CIFAR-۱۰ and CIFAR-۱۰۰](#)
3. [IMAGENET](#)
4. [Tiny Images](#) ۸۰ Million tiny images
5. [Flickr Data](#) ۱۰۰ Million Yahoo dataset
6. [Berkeley Segmentation Dataset](#) ۵۰۰
7. [UC Irvine Machine Learning Repository](#)
8. [Flickr](#) ۸k
9. [Flickr](#) ۳۰k
10. [Microsoft COCO](#)
11. [VQA](#)
12. [Image QA](#)
13. [AT&T Laboratories Cambridge face database](#)
14. [AVHRR Pathfinder](#)
15. [Air Freight](#) – The Air Freight data set is a ray-traced image sequence along with ground truth segmentation based on textural characteristics. (۴۵۵ images + GT, each ۱۶۰×۱۲۰ pixels). (Formats: PNG)
16. [Amsterdam Library of Object Images](#) – ALOI is a color image collection of one-thousand small objects, recorded for scientific purposes. In order to capture the sensory variation in object recordings, we systematically varied viewing angle, illumination angle, and illumination color for each object, and additionally captured wide-baseline stereo images. We recorded over a hundred images of each object, yielding a total of ۱۱۰,۲۵۰ images for the collection. (Formats: png)

Annotated face, hand, cardiac & meat images – Most images & annotations are supplemented by various ASM/AAM analyses using the AAM-API. (Formats: bmp, asf)	.17
Image Analysis and Computer Graphics	.18
Brown University Stimuli – A variety of datasets including geons, objects, and “greebles”. Good for testing recognition algorithms. (Formats: pict)	.19
CAVIAR video sequences of mall and public space behavior – ۹۰K video frames in ۹۰ sequences of various human activities, with XML ground truth of detection and behavior classification (Formats: MPEG۲ & JPEG)	.20
Machine Vision Unit	.21
CCITT Fax standard images – ۸ images (Formats: gif)	.22
CMU CIL’s Stereo Data with Ground Truth – ۳ sets of ۱۱ images, including color tiff images with spectroradiometry (Formats: gif, tiff)	.23
CMU PIE Database – A database of ۴۱,۳۶۸ face images of ۶۸ people captured under ۱۳ poses, ۴۳ illuminations conditions, and with ۴ different expressions	.24
CMU VASC Image Database – Images, sequences, stereo pairs (thousands of images) (Formats: Sun Rasterimage)	.25
Caltech Image Database – about ۲۰ images – mostly top-down views of small objects and toys. (Formats: GIF)	.26
Columbia-Utrecht Reflectance and Texture Database – Texture and reflectance measurements for over ۶۰ samples of ۳D texture, observed with over ۲۰۰ different combinations of viewing and illumination directions. (Formats: bmp)	.27
Computational Colour Constancy Data – A dataset oriented towards computational color constancy, but useful for computer vision in general. It includes synthetic data, camera sensor data, and over ۷۰۰ images. (Formats: tiff)	.28
Computational Vision Lab	.29
Content-based image retrieval database – ۱۱ sets of color images for testing algorithms for content-based retrieval. Most sets have a description file with names of objects in each image. (Formats: jpg)	.30
Efficient Content-based Retrieval Group	.31
Densely Sampled View Spheres – Densely sampled view spheres – upper half of the view sphere of two toy objects with ۲۵۰۰ images each. (Formats: tiff)	.32
Computer Science VII (Graphical Systems)	.33

Digital Embryos – Digital embryos are novel objects which may be used to develop and test object recognition systems. They have an organic appearance. (Formats: various formats are available on request)	.34
University of Minnesota Vision Lab	.35
El Salvador Atlas of Gastrointestinal VideoEndoscopy – Images and Videos of his-res of studies taken from Gastrointestinal Video endoscopy. (Formats: jpg, mpg, gif)	.36
FG-NET Facial Aging Database – Database contains ۱۰۰۲ face images showing subjects at different ages. (Formats: jpg)	.37
FVC۲۰۰۰ Fingerprint Databases – FVC۲۰۰۰ is the First International Competition for Fingerprint Verification Algorithms. Four fingerprint databases constitute the FVC۲۰۰۰ benchmark (۳۵۲۰ fingerprints in all)	.38
Biometric Systems Lab – University of Bologna	.39
Face and Gesture images and image sequences – Several image datasets of faces and gestures that are ground truth annotated for benchmarking	.40
German Fingerspelling Database – The database contains ۲۵ gestures and consists of ۱۴۰۰ image sequences that contain gestures of ۲۰ different persons recorded under non-uniform daylight lighting conditions. (Formats: mpg, jpg)	.41
Language Processing and Pattern Recognition	.42
Groningen Natural Image Database – ۴۰۰۰+ ۱۵۳۶×۱۰۲۴ (۱۶ bit) calibrated outdoor images (Formats: homebrew)	.43
ICG Testhouse sequence – ۲ turntable sequences from different viewing heights, ۳۶ images each, resolution ۱۰۰۰×۷۵۰, color (Formats: PPM)	.44
Institute of Computer Graphics and Vision	.45
IEN Image Library – ۱۰۰۰+ images, mostly outdoor sequences (Formats: raw, ppm)	.46
INRIA's Syntim images database – ۱۵ color image of simple objects (Formats: gif)	.47
INRIA	.48
INRIA's Syntim stereo databases – ۳۴ calibrated color stereo pairs (Formats: gif)	.49
Image Analysis Laboratory – Images obtained from a variety of imaging modalities — raw CFA images, range images and a host of “medical images”. (Formats: homebrew)	.50
Image Analysis Laboratory	.51

- [Image Database](#) - An image database including some textures .52
- [JAFPE Facial Expression Image Database](#) - The JAFPE database consists of ۲۱۳ images of Japanese female subjects posing ۶ basic facial expressions as well as a neutral pose. Ratings on emotion adjectives are also available, free of charge, for research purposes. (Formats: TIFF Grayscale images.) .53
- [ATR Research, Kyoto, Japan](#) .54
- JISCT Stereo Evaluation - ۴۴ image pairs. These data have been used in an evaluation of stereo analysis, as described in the April ۱۹۹۳ ARPA Image Understanding Workshop paper "The JISCT Stereo Evaluation" by R.C.Bolles, H.H.Baker, and M.J.Hannah, ۲۶۳-۲۷۴ (Formats: SSI) .۵۵
- [MIT Vision Texture](#) - Image archive (۱۰۰+ images) (Formats: ppm) .56
- MIT face images and more - hundreds of images (Formats: homebrew) .۵۷
- [Machine Vision](#) - Images from the textbook by Jain, Kasturi, Schunck (۲۰+ images) (Formats: GIF TIFF) .58
- [Mammography Image Databases](#) - ۱۰۰ or more images of mammograms with ground truth. Additional images available by request, and links to several other mammography databases are provided. (Formats: homebrew) .59
- <ftp://ftp.cps.msu.edu/pub/prip> - many images (Formats: unknown) .60
- [Middlebury Stereo Data Sets with Ground Truth](#) - Six multi-frame stereo data sets of scenes containing planar regions. Each data set contains ۹ color images and subpixel-accuracy ground-truth data. (Formats: ppm) .61
- [Middlebury Stereo Vision Research Page](#) - Middlebury College .62
- [Modis Airborne simulator, Gallery and data set](#) - High Altitude Imagery from around the world for environmental modeling in support of NASA EOS program (Formats: JPG and HDF) .63
- NIST Fingerprint and handwriting - datasets - thousands of images (Formats: unknown) .۶۴
- NIST Fingerprint data - compressed multipart uuencoded tar file .۶۵
- [NLM HyperDoc Visible Human Project](#) - Color, CAT and MRI image samples - over ۲۰ images (Formats: jpeg) .66
- [National Design Repository](#) - Over ۵۵,۰۰۰ ۳D CAD and solid models of (mostly) mechanical/machined engineerign designs. (Formats: gif,vrml,wrl,stp,sat) .67
- [Geometric & Intelligent Computing Laboratory](#) .68

- [OSU \(MSU\) ۳D Object Model Database](#) – several sets of ۳D object models collected over several years to use in object recognition research (Formats: homebrew, vrml) .69
- [OSU \(MSU/WSU\) Range Image Database](#) – Hundreds of real and synthetic images (Formats: gif, homebrew) .70
- [OSU/SAMPL Database: Range Images, ۳D Models, Stills, Motion Sequences](#) – Over ۱۰۰۰ range images, ۳D object models, still images and motion sequences (Formats: gif, ppm, vrml, homebrew) .71
- [Signal Analysis and Machine Perception Laboratory](#) .72
- [Otago Optical Flow Evaluation Sequences](#) – Synthetic and real sequences with machine-readable ground truth optical flow fields, plus tools to generate ground truth for new sequences. (Formats: ppm,tif,homebrew) .73
- [Vision Research Group](#) .74
- <ftp://ftp.limsi.fr/pub/quenot/opflow/testdata/piv/> – Real and synthetic image sequences used for testing a Particle Image Velocimetry application. These images may be used for the test of optical flow and image matching algorithms. (Formats: pgm (raw)) .75
- [LIMSI-CNRS/CHM/IMM/vision](#) .76
- [LIMSI-CNRS](#) .77
- [Photometric ۳D Surface Texture Database](#) – This is the first ۳D texture database which provides both full real surface rotations and registered photometric stereo data (۳ textures, ۱۶۸۰ images). (Formats: TIFF) .78
- [SEQUENCES FOR OPTICAL FLOW ANALYSIS \(SOFA\)](#) – ۹ synthetic sequences designed for testing motion analysis applications, including full ground truth of motion and camera parameters. (Formats: gif) .79
- [Computer Vision Group](#) .80
- [Sequences for Flow Based Reconstruction](#) – synthetic sequence for testing structure from motion algorithms (Formats: pgm) .81
- [Stereo Images with Ground Truth Disparity and Occlusion](#) – a small set of synthetic images of a hallway with varying amounts of noise added. Use these images to benchmark your stereo algorithm. (Formats: raw, viff (khoros), or tiff) .82
- [Stuttgart Range Image Database](#) – A collection of synthetic range images taken from high-resolution polygonal models available on the web (Formats: homebrew) .83
- [Department Image Understanding](#) .84

- [The AR Face Database](#) – Contains over ۴,۰۰۰ color images corresponding to ۱۲۶ people's faces (۷۰ men and ۵۶ women). Frontal views with variations in facial expressions, illumination, and occlusions. (Formats: RAW (RGB ۲۴-bit)) .85
- [Purdue Robot Vision Lab](#) .86
- [The MIT-CSAIL Database of Objects and Scenes](#) – Database for testing multiclass object detection and scene recognition algorithms. Over ۷۲,۰۰۰ images with ۲۸۷۳ annotated frames. More than ۵۰ annotated object classes. (Formats: jpg) .87
- [The RVL SPEC-DB \(SPECularity DataBase\)](#) – A collection of over ۳۰۰ real images of ۱۰۰ objects taken under three different illuminaiton conditions (Diffuse/Ambient/Directed). — Use these images to test algorithms for detecting and compensating specular highlights in color images. (Formats: TIFF) .88
- [Robot Vision Laboratory](#) .89
- [The XmrvtS database](#) – The XM۲VTSDb contains four digital recordings of ۲۹۵ people taken over a period of four months. This database contains both image and video data .of faces .90
- [Centre for Vision, Speech and Signal Processing](#) .91
- [Traffic Image Sequences and 'Marbled Block' Sequence](#) – thousands of frames of digitized traffic image sequences as well as the 'Marbled Block' sequence (grayscale images) (Formats: GIF) .92
- [IAKS/KOGS](#) .93
- U Bern Face images – hundreds of images (Formats: Sun rasterfile) .۹۴
- U Michigan textures (Formats: compressed raw) .۹۵
- [U Oulu wood and knots database](#) – Includes classifications – ۱۰۰۰+ color images (Formats: ppm) .96
- [UCID – an Uncompressed Colour Image Database](#) – a benchmark database for image retrieval with predefined ground truth. (Formats: tiff) .97
- [UMass Vision Image Archive](#) – Large image database with aerial, space, stereo, medical images and more. (Formats: homebrew) .98
- UNC's ۳D image database – many images (Formats: GIF) .۹۹
- [USF Range Image Data with Segmentation Ground Truth](#) – ۸۰ image sets (Formats: Sun rasterimage) .100

- [University of Oulu Physics-based Face Database](#) – contains color images of faces under different illuminants and camera calibration conditions as well as skin spectral reflectance measurements of each person .101
- [Machine Vision and Media Processing Unit](#) .102
- [University of Oulu Texture Database](#) – Database of ۳۲۰ surface textures, each captured under three illuminants, six spatial resolutions and nine rotation angles. A set of test suites is also provided so that texture segmentation, classification, and retrieval algorithms can be tested in a standard manner. (Formats: bmp, ras, xv) .103
- [Machine Vision Group](#) .104
- Usenix face database – Thousands of face images from many different sites (circa ۹۹۴) .۱۰۵
- [View Sphere Database](#) – Images of ۸ objects seen from many different view points. The view sphere is sampled using a geodesic with ۱۷۲ images/sphere. Two sets for training and testing are available. (Formats: ppm) .106
- [PRIMA, GRAVIR](#) .107
- Vision-list Imagery Archive – Many images, many formats .۱۰۸
- [Wiry Object Recognition Database](#) – Thousands of images of a cart, ladder, stool, bicycle, chairs, and cluttered scenes with ground truth labelings of edges and regions. (Formats: jpg) .109
- [۳ D Vision Group](#) .110
- [Yale Face Database](#) – ۱۶۵ images (۱۵ individuals) with different lighting, expression, and occlusion configurations .111
- [Yale Face Database B](#) – ۵۷۶۰ single light source images of ۱۰ subjects each seen under ۵۷۶ viewing conditions (۹ poses x ۶۴ illumination conditions). (Formats: PGM) .112
- [Center for Computational Vision and Control](#) .113
- [DeepMind QA Corpus](#) – Textual QA corpus from CNN and DailyMail. More than ۳۰۰K documents in total. [Paper](#) for reference .114